-24-

CLAIMS

1. A system for generating an operational assessment of a cache memory in a digital data processing system for respective cache memory sizes comprising:

    A.   an operational statistics gathering element for gathering operational statistics over a time interval, including a file information retrieval activity value and a extent of activity value for each file accessed during the time interval;

    B.   a cache miss prediction element for generating a cache miss prediction value in response to the operational statistics gathered by the operational statistics gathering element and a cache memory size value; and

    C.   a cache memory size adjustment element for adjusting the cache memory size in response to the cache memory size value generated by the cache miss prediction element for a selected one of said cache miss prediction values.

2. A system as defined in claim 1 in which the cache miss prediction element generates the cache miss prediction value based on a particular one of a plurality of cache memory management methodologies.
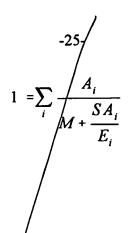
3. A system as defined in claim 2 in which one of said cache memory management methodologies is a FIFO (first-in/first-out) methodology, the cache miss prediction element generating the cache miss prediction value in accordance with:

$$1 = \sum_i \frac{A_i}{M + \dfrac{SA_i}{E_i}}$$

-25-

4

5   where "M" represents the cache miss prediction value, "S" represents the selected cache memory size

6   value, "A_i" represents the file retrieval activity value for a file "i," and "E_i" represents the extent of

7   activity value for the file "i."

1   4. A system as defined in claim 3 in which the cache miss prediction element determines the cache

2   miss prediction value "M" using a binary search methodology over the interval

3   $$0 \le M \le \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.

1   5. A system as defined in claim 2 in which one of said cache memory management methodologies is

2   a FIFO (first-in/first-out) methodology, the operational statistics gathering element further gathering

3   a packet re-reference value indicating a number of times a portion of a file, identified as a packet, is

4   referenced during the time interval, the cache miss prediction element generating the cache miss

5   prediction value in accordance with:

6

$$1 = \sum_i \frac{\dfrac{A_i}{R_i}(R_i - 1 - H(i,S))}{M + \dfrac{S\left(\dfrac{A_i}{R_i}\right)}{E_i}}$$

7   where "M" represents the cache miss/prediction value, "S" represents the selected cache memory size

8   value, "A_i" represents the file retrieval activity value for a file "i," "E_i" represents the extent of activity

9   value for the file "i," and "R_i" represents the packet re-reference value for file "i."

1   6. A system as defined in claim 5 in which the cache miss prediction element determines the cache

2   miss prediction value "M" using a binary search methodology over the interval

3
$$0 \le M \le \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.

1   7. A system as defined in claim 2 in which one of said cache memory management methodologies is

2   an LRU (least-recently used) methodology, the operational statistics gathering element further

3   gathering a packet re-reference value indicating a number of times a portion of a file, identified as a

4   packet, is referenced during the time interval, the cache miss prediction element generating the cache

5   mis prediction value in accordance with:

6

$$1 = \sum_i \frac{\left(\dfrac{A_i T_i}{R_i}\right) M + \left(\dfrac{A_i SI}{R_i}\right)}{\left(SI + \dfrac{SA_i T_i}{E_i R_i}\right) M + \left(\dfrac{A_i I}{E_i R_i}\right) S^2}$$

7   where "M" represents the cache miss prediction value, "S" represents the selected cache memory size

8   value, "A_i" represents the file retrieval activity value for a file "i," "E_i" represents the extent of activity

9   value for the file "i," "R_i" represents the packet re-reference value for file "i," and "I" represents the

10   duration of the time interval.

1   8. A system as defined in claim 7 in which the cache miss prediction element determines the cache

2   miss prediction value "M" using a binary search methodology over the interval

3

$$0 \le M \le \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.

1 9. A method for generating an operational assessment of a cache memory in a digital data processing

2 system for respective cache memory sizes comprising the steps of:

3 A. gathering operational statistics over a time interval, including a file information retrieval

4 activity value and a extent of activity value for each file accessed during the time interval;

5 B. generating a cache miss prediction value in response to the operational statistics gathered

6 during the operational statistics gathering step, and a cache memory size value; and

7 C. adjusting the cache memory size in response to the cache memory size value generated during

8 the cache miss prediction step for a selected one of said cache miss prediction values.

1 10. A method as defined in claim 9 in which during the cache miss prediction step the cache miss

2 prediction value based on a particular one of a plurality of cache memory management

3 methodologies.

1 11. A method as defined in claim 10 in which one of said cache memory management methodologies

2 is a FIFO (first-in/first-out) methodology, during the cache miss prediction step the cache miss

3 prediction value being generated in accordance with:

4

$$1 = \sum_i \frac{A_i}{M + \dfrac{SA_i}{E_i}}$$

5   where "M" represents the cache miss prediction value, "S" represents the selected cache memory size

6   value, "A_i" represents the file retrieval activity value for a file "i," and "E_i" represents the extent of

7   activity value for the file "i."

1   12. A method as defined in claim 11 in which, during the cache miss prediction step, the cache miss

2   prediction value "M" being generated using a binary search methodology over the interval

$$0 \le M \le \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.

13. A method as defined in claim 10 in which one of said cache memory management methodologies

2   is a FIFO (first-in/first-out) methodology, the operational statistics gathering element further

3   gathering a packet re-reference value indicating a number of times a portion of a file, identified as a

4   packet, is referenced during the time interval, during the cache miss prediction step the cache miss

5   prediction value being generated in accordance with:

$$1 = \sum_i \frac{\frac{A_i}{R_i}(R_i - 1 - H(i,S))}{M + \frac{S\left(\frac{A_i}{R_i}\right)}{E_i}}$$

7   where "M" represents the cache miss prediction value, "S" represents the selected cache memory size

8   value, "A_i" represents the file retrieval activity value for a file "i," "E_i" represents the extent of activity

9   value for the file "i," and "R_i" represents the packet re-reference value for file "i."

1   14. A method as defined in claim 13 in which, during cache miss prediction step, the cache miss

2   prediction value "M" being generated using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.

15. A method as defined in claim 10 in which one of said cache memory management methodologies

is an LRU (least-recently used) methodology, during the operational statistics gathering step a packet

3   re-reference value being further gathered indicating a number of times a portion of a file, identified

4   as a packet, is referenced during the time interval, the cache miss prediction element generating the

5   cache mis prediction value in accordance with:

$$1 = \sum_i \frac{\left(\frac{A_i T_i}{R_i}\right) M + \left(\frac{A_i SI}{R_i}\right)}{\left(SI + \frac{S A_i T_i}{E_i R_i}\right) M + \left(\frac{A_i I}{E_i R_i}\right) S^2}$$

7   where "M" represents the cache miss prediction value, "S" represents the selected cache memory size

8   value, "A_i" represents the file retrieval activity value for a file "i," "E_i" represents the extent of activity

9   value for the file "i," "R_i" represents the packet re-reference value for file "i," and "I" represents the

10   duration of the time interval.

1   16. A method as defined in claim 15 in which during the cache miss prediction step the cache miss

2   prediction value "M" being generated using a binary search methodology over the interval

$$0 \le M \le \sum_i A_i = A$$

4   where "A" represents the total activity over the time interval.